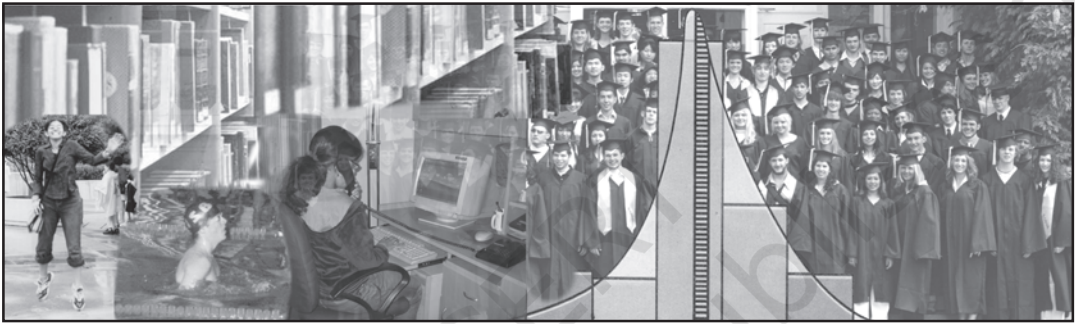
11099CH03

अध्याय
3

आँकड़ों का संगठन

इस अध्याय को पढ़ने के बाद आप इस योग्य होंगे कि:

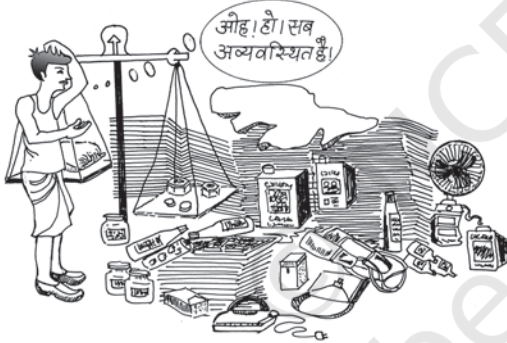
- *आगे के सांख्यिकीय विश्लेषण के लिए आँकड़ों का वर्गीकरण कर सकें;*
- *मात्रात्मक एवं गुणात्मक वर्गीकरण के बीच अंतर कर सकें;*
- *बारंबारता वितरण सारणी तैयार कर सकें;*
- *वर्गों के निर्माण की तकनीक जान सकें;*
- *मिलान-चिह्न की विधि से परिचित हो सकें;*
- *एकचर तथा द्विचर बारंबारता वितरण के बीच अंतर कर सकें।*

1. प्रस्तावना

पिछले अध्याय में आपने पढ़ा कि आँकड़ों का संग्रहण कैसे करते हैं। साथ ही, आप जनगणना एवं प्रतिचयन के बीच अंतर को भी जान चुके हैं। इस अध्याय में आप यह सीखेंगे कि जो आँकड़ें आपने संगृहीत किए थे, उन्हें कैसे वर्गीकृत करते हैं। अपरिष्कृत आँकड़ों को वर्गीकृत करने का उद्देश्य उन्हें व्यवस्थित करना है, ताकि उन्हें आसानी से आगे के सांख्यिकीय विशलेषण के योग्य बनाया जा सके।

क्या आपने कभी स्थानीय कबाड़ी वाले या रद्दी सामान खरीदने वाले को देखा है, जिसे आप अपना पुराना अखबार, टूटे-फूटे घरेलू सामान, खाली-काँच की बोतलें, प्लास्टिक आदि बेचते हैं। वह आपसे इन चीजों को खरीदता है और उन लोगों को बेच देता है जो इनका पुनः चक्रण करते हैं। लेकिन अपनी दुकान में अधिक कबाड़ के इकट्ठे होने से उसे अपना व्यापार चलाने में मुश्किल हो सकती है, अगर वह इन्हें उचित ढंग से व्यवस्थित न करे। वह इस स्थिति को सरल बनाने के लिए विभिन्न कबाड़ों को उपयुक्त समूह में रखता है, अर्थात् उन्हें *वर्गीकृत* करता है। वह पुराने अखबारों को एक साथ रस्सी से बाँध कर रखता है। इसके बाद सभी खाली काँच की बोतलों को एक बोरे

में रखता है। वह धातु के सामानों का एक ढेर अपनी दुकान के एक कोने में लगाता है और फिर उनको 'लोहा', 'पीतल', 'ताँबा', 'एल्यूमिनियम' आदि वर्गों में छाँट कर रखता है। इस प्रकार से वह अपने कबाड़ को भिन्न वर्गों – 'अखबार', 'प्लास्टिक', 'काँच', 'धातु' आदि में विभाजित कर उन्हें व्यवस्थित करता है। जब एक बार उसका सारा कबाड़ व्यवस्थित एवं वर्गीकृत हो जाता है, तब खरीददार की माँग पर, उसे सामग्री विशेष को खोजकर देने में आसानी हो जाती है।

ठीक इसी प्रकार से, जब आप अपने विद्यालय की पुस्तकों को एक विशेष क्रम में रखते हैं, तो उनको संभालना आसान हो जाता है। आप उन्हें विषयों के अनुसार वर्गीकृत कर सकते हैं, जहाँ प्रत्येक विषय एक

समूह या वर्ग बन जाता है। उदाहरणार्थ, जब आपको इतिहास की कोई विशेष पुस्तक की आवश्यकता पड़ती है तो आप को केवल यह करना है कि 'इतिहास' समूह में उस पुस्तक को खोजें। अन्यथा आप को अपनी यह विशेष पुस्तक सारी पुस्तकों के ढेर में खोजनी पड़ेगी।

यद्यपि पदार्थों अथवा वस्तुओं का वर्गीकरण बहुमूल्य श्रम और समय को बचाता है, इसे मनमाने तरीके से नहीं किया जाता है। कबाड़ी वाले ने अपने कबाड़ को इस तरह से समूहों में रखा कि प्रत्येक समूह में एक ही प्रकार की चीजें हों। उदाहरण के लिए, उसने 'काँच' के समूह में खाली काँच की बोतलें, टूटे खिड़की के काँच तथा टूटे दर्पण आदि रखे। ठीक इसी तरह से जब आपने अपनी इतिहास की पुस्तक को 'इतिहास' समूह में वर्गीकृत किया, तो आप उसमें अन्य विषयों की पुस्तकें नहीं रखेंगे। अन्यथा समूह–गठन का पूरा उद्देश्य ही निरर्थक हो जाएगा। *इसलिए, वर्गीकरण का तात्पर्य एक वस्तुओं को समूह या वर्गों में किसी खास आधार पर वर्गीकृत या व्यवस्थित करने से है।*

क्रियात्मक गतिविधि

- अपने स्थानीय डाकघर जायें और देखें कि पत्रों कि छँटाई कैसे की जाती है। क्या आप जानते हैं कि पत्र में पिन कोड का क्या अर्थ है। अपने डाकिए से पूछें।

2. अपरिष्कृत आँकड़े

कबाड़ीवाले के कबाड़ की भाँति, अवर्गीकृत आँकड़े अथवा *अपरिष्कृत आँकड़े* भी अत्यधिक अव्यवस्थित होते हैं। ये प्राय: अति विशाल होते हैं, जिन्हें संभालना कठिन होता है। इनसे सार्थक निष्कर्ष निकालना श्रमसाध्य कार्य है, क्योंकि सांख्यिकीय विधियों का इन पर सरलता से प्रयोग नहीं किया जा सकता। इसलिए इस प्रकार के आँकड़ों का उचित संगठन तथा प्रस्तुतीकरण आवश्यक होता है, ताकि व्यवस्थित रूप से साँख्यिकीय विश्लेषण किया जा सके। अत: आँकड़ों के संग्रह के पश्चात् अगला चरण उन्हें संगठित कर वर्गीकृत रूप में प्रस्तुत करना है।

मान लीजिए, कि आप गणित में छात्रों की प्रगति जानना चाहते हैं और आपने अपने स्कूल के 100 छात्रों के गणित के अंकों के आँकड़े एकत्रित कर लिये हैं। अगर आप इन्हें एक सारणी में प्रस्तुत करते हैं तो वे संभवत: सारणी 3.1 जैसे प्रतीत हो सकते हैं।

सारणी 3.1

किसी परीक्षा में 100 छात्रों द्वारा गणित में प्राप्त अंक

47	45	10	60	51	56	66	100	49	40
60	59	56	55	62	48	59	55	51	41
42	69	64	66	50	59	57	65	62	50
64	30	37	75	17	56	20	14	55	90
62	51	55	14	25	34	90	49	56	54
70	47	49	82	40	82	60	85	65	66
49	44	64	69	70	48	12	28	55	65
49	40	25	41	71	80	0	56	14	22
66	53	46	70	43	61	59	12	30	35
45	44	57	76	82	39	32	14	90	25

या फिर आप अपने पड़ोस के 50 परिवारों से, भोजन पर उनके मासिक व्यय के आँकड़ों का संग्रह यह जानने के लिए करते हैं कि भोजन पर उनका औसत व्यय कितना है। इस मामले में संगृहीत आँकड़ों को जब आप सारणी में प्रस्तुत करते हैं, तो वे सारणी 3.2 की तरह दिख सकते हैं। सारणी 3.1 तथा सारणी 3.2, दोनों ही आँकड़े अपरिष्कृत अथवा अवर्गीकृत हैं। दोनों ही सारणियों में संख्याओं को किसी भी क्रम में

व्यवस्थित नहीं किया गया है। अब अगर आपसे यह पूछा जाए कि सारणी 3.1 में गणित में सर्वोच्च अंक कितने हैं, तब आपको 100 छात्रों के अंकों को आरोही या अवरोही क्रम में व्यवस्थित करना होगा। यह एक बेहद थका देने वाला काम है। यदि आपको 100 छात्रों के स्थान पर 1000 छात्रों के अंक संभालने हों तो यह और भी अधिक थकानेवाला होगा।

सारणी 3.2

खाद्य पर 50 परिवारों के मासिक पारिवारिक व्यय (रु में)

1904	1559	3473	1735	2760
2041	1612	1753	1855	4439
5090	1085	1823	2346	1523
1211	1360	1110	2152	1183
1218	1315	1105	2628	2712
4248	1812	1264	1183	1171
1007	1180	1953	1137	2048
2025	1583	1324	2621	3676
1397	1832	1962	2177	2575
1293	1365	1146	3222	1396

ठीक इसी प्रकार से, सारणी 3.2 में आपके लिए काफी मुश्किल होगा कि 50 परिवारों के खाने पर मासिक व्यय के औसत को पता कर सकें। यही कठिनाई तब कई गुना बढ़ जाएगी यदि यह संख्या बहुत बड़ी हो, जैसे 5000 परिवार। ठीक कबाड़ीवाले की भाँति ही (जब कबाड़ का ढेर बहुत बड़ा और अव्यवस्थित हो तो उसे एक विशेष वस्तु को ढूँढ़ने में बहुत कठिनाई होती है) आपकी भी स्थिति होगी, यदि अपरिष्कृत आँकड़ों का भंडार बहुत बड़ा हो और आप उससे कोई सूचना प्राप्त करना चाहें। इसलिए, संक्षेप में अवर्गीकृत विशाल आँकड़ों से कोई सूचना प्राप्त करना एक बेहद थका देने वाला एवं उबाऊ काम है।

वर्गीकरण के द्वारा अपरिष्कृत आँकड़ों को संक्षिप्त एवं बोधगम्य बनाया जाता है। जब एक प्रकार की विशेषताओं वाले तथ्यों को एक ही वर्ग में रखा जाता है तो वे बिना किसी कठिनाई के ढूँढ़ने, तुलना करने तथा निष्कर्ष निकालने योग्य हो जाते हैं। आपने अध्याय 2 में पढ़ा है कि प्रति दस साल बाद भारत सरकार जनसंख्या की गणना कराती है। सन् 2001 की

जनगणना में लगभग 20 करोड़ लोगों से संपर्क किया गया। जनगणना के अपरिष्कृत आँकड़े बहुत विशाल एवं विखंडित होते हैं। उन से कोई भी अर्थपूर्ण निष्कर्ष निकालना असंभव कार्य लगता है। लेकिन जनगणना के यही आँकड़े जब लिंग, शिक्षा, वैवाहिक स्थिति, पेशे आदि के अनुसार वर्गीकृत किये जाते हैं तब भारत की जनसंख्या की प्रकृति एवं संरचना आसानी से समझ में आ जाती है।

अपरिष्कृत आँकड़े चरों के प्रेक्षणों से बने होते हैं। सारणी 3.1 तथा 3.2 में दिए गए अपरिष्कृत आँकड़े विशेष या चर समूह पर किए गए प्रेक्षणों को प्रदर्शित करते हैं। उदाहरण के लिए सारणी 3.1 को देखें जिसमें 100 छात्रों द्वारा गणित में प्राप्त किए गए अंकों को दर्शाया गया है। इन अंकों से हम कैसे अर्थ निकाल सकते हैं? गणित के शिक्षक इन अंकों को देखकर सोच रहे होंगे – मेरे छात्रों ने कैसा किया? कितने असफल रहे? आँकड़ों का वर्गीकरण हमारे उद्देश्यों पर निर्भर करता है इस स्थिति में, शिक्षक गहनतापूर्वक समझने की कोशिश करेंगे – छात्रों ने कैसा किया? संभवतया वह बारंबारता वितरण बनाने का चयन करे। इस पर अगले भाग में विवेचना की जायेगी।

क्रियात्मक गतिविधि

- आप अपने परिवार के एक वर्ष के साप्ताहिक व्यय के आँकड़े संगृहीत कीजिए और उसे एक सारणी में व्यवस्थित कीजिए। देखिए कि उसमें कितने प्रेक्षण हैं। आँकड़ों को मासिक आधार पर व्यवस्थित कीजिए और देखिए कि अब कितने प्रेक्षण हैं।

3. आँकड़ों का वर्गीकरण

किसी वर्गीकरण के वर्ग या समूह कई तरीकों से बनाए जा सकते हैं। आप अपनी पुस्तकों को विषयों–'इतिहास', 'भूगोल', 'गणित', 'विज्ञान' आदि में वर्गीकृत करने के स्थान पर इन्हें वर्णमाला के क्रम में लेखकों के आधार पर वर्गीकृत कर सकते हैं। अथवा, आप इन्हें प्रकाशन– वर्ष के आधार पर भी वर्गीकृत कर सकते हैं। आप उन्हें किस प्रकार से वर्गीकृत करना चाहते हैं, यह आपकी आवश्यकता पर निर्भर करेगा।

ठीक इसी प्रकार से, अपरिष्कृत आँकड़ों को भी विभिन्न तरीकों से वर्गीकृत किया जा सकता है जो आपके अध्ययन के उद्देश्य पर निर्भर करता है। उन्हें समय के अनुसार समूहित किया जा सकता है। इस प्रकार के वर्गीकरण को **कालानुक्रमिक वर्गीकरण** कहते हैं। इस प्रकार के वर्गीकरण में, आँकड़ों को समय के संदर्भ–जैसे वर्ष, तिमाही, मासिक या साप्ताहिक आदि के रूप में, आरोही या अवरोही क्रम में वर्गीकृत किया जा सकता है। निम्नलिखित उदाहरण वर्षों के आधार पर भारत की जनसंख्या के वर्गीकरण को दिखाता है। चर 'जनसंख्या' एक **काल–श्रेणी** है, क्योंकि इसमें विभिन्न वर्षों के मानों की एक श्रेणी चित्रित की गई है।

उदाहरण 1

भारत की जनसंख्या (करोड़ में)

वर्ष	*जनसंख्या (करोड़ में)*
1951	35.7
1961	43.8
1971	54.6
1981	68.4
1991	81.8
2001	102.7
2011	121.0

स्थानिक वर्गीकरण के अंतर्गत आँकड़ों का वर्गीकरण भौगोलिक स्थितियों जैसे कि देश, राज्य, शहर, जिला, कस्बा आदि के संदर्भानुसार होता है। उदाहरण 2 में विभिन्न देशों में गेहूँ की उपज दिखाई गई है।

उदाहरण 2

विभिन्न देशों में गेहूँ की उपज (2013)

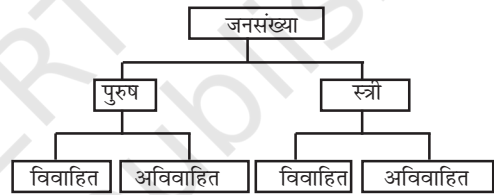
देश	*गेहूँ की उपज (किग्रा/एकड़)*
कनाडा	3594
चीन	5055
फ्रांस	7254
जर्मनी	7998
भारत	3154
पाकिस्तान	2787

स्रोतः कृषि आँकड़े, भारत सरकार, 2015

क्रियात्मक गतिविधियाँ

- उदाहरण 1 में, उस वर्ष को बताएँ जिसमें भारत की जनसंख्या न्यूनतम और अधिकतम है।
- उदाहरण 2 में, उस देश का पता लगाइये, जिसकी गेहूँ की उपज भारत से थोड़ी अधिक है। यह प्रतिशत में कितनी होगी?
- उदाहरण दो में दिए गए देशों को गेहूँ की उपज के आरोही क्रम में रखिये। ठीक यही अभ्यास उपज को अवरोही क्रम में रखते हुए कीजिए।

कई बार आपका सामना ऐसी विशेषताओं से होता है, जिन्हें मात्रात्मक रूप में व्यक्त नहीं किया जा सकता है। इस प्रकार की विशेषताओं को *'गुण'* कहते हैं। उदाहरण के लिए-राष्ट्रीयता, साक्षरता, धर्म, लिंग, वैवाहिक स्थिति आदि। इन्हें मापा नहीं जा सकता है। इन गुणों को *गुणात्मक* विशेषता की उपस्थिति या अनुपस्थिति के आधार पर वर्गीकृत कर सकते हैं। विशेषताओं पर आधारित आँकड़ों के ऐसे वर्गीकरण को *गुणात्मक वर्गीकरण* कहा जाता है। निम्नलिखित उदाहरण में हम किसी देश की जनसंख्या को गुणात्मक चर 'लिंग' के आधार पर समूहित किया हुआ पाते हैं। इसमें प्रेक्षण स्त्री या पुरुष हो सकता है। इन दो विशेषताओं को आगे वैवाहिक स्थिति के आधार पर वर्गीकृत किया जा सकता है, जैसा कि नीचे दिया गया है

उदाहरण 3

प्रथम चरण में यह वर्गीकरण पहले विशेषता की उपस्थिति या अनुपस्थिति पर आधारित है जैसे कि 'पुरुष' या 'पुरुष नहीं' (स्त्री) है। दूसरे चरण में, प्रत्येक वर्ग 'स्त्री' या 'पुरुष' आगे दूसरी विशेषता की उपस्थिति या अनुपस्थिति के आधार पर विभाजित है, जैसे विवाहित या अविवाहित। ऊँचाई, भार, आयु, आय, छात्रों के अंक आदि विशेषताओं की प्रकृति मात्रात्मक है। जब ऐसी विशिष्टताओं के संगृहीत आँकड़ों को वर्गो में समूहित किया जाता है तो यह वर्गीकरण *मात्रात्मक वर्गीकरण* कहलाता है।

क्रियात्मक गतिविधियाँ

- आस-पास की वस्तुओं को सजीव या निर्जीव के रूप में समूहित किया जा सकता है। क्या यह मात्रात्मक वर्गीकरण है?

उदाहरण 4

100 छात्रों के गणित के प्राप्तांकों का बारंबारता वितरण

अंक	*बारंबारता*
0–10	1
10–20	8
20–30	6
30–40	7
40–50	21
50–60	23
60–70	19
70–80	6
80–90	5
90–100	4
योग	*100*

उदाहरण 4 में 100 छात्रों के गणित के प्राप्तांकों का मात्रात्मक वर्गीकरण दिखाया गया है, जिन्हें सारणी 3.1 में बारंबारता वितरण के रूप में दिया गया है।

क्रियात्मक गतिविधियाँ

- उदाहरण 4 की बारंबारता के मानों को कुल बारंबारता के अनुपात में या प्रतिशत में प्रकट कीजिए। ध्यान रहे कि इस प्रकार से प्रकट की गई बारंबारता को सापेक्षिक बारंबारता के रूप में जाना जाता है।
- उदाहरण 4 में किस वर्ग के अंतर्गत आँकड़ों का अधिकतम संकेंद्रण है? इसे कुल प्रेक्षणों के प्रतिशत के रूप में प्रकट कीजिए। किस वर्ग में आँकड़ों का न्यूनतम संकेंद्रण है?

4. चर : संतत और विविक्त

चर की सरल परिभाषा, जिसका आपने पिछले अध्याय में अध्ययन किया था, यह नहीं बतलाती कि यह कैसे परिवर्तित होता है। चरों में अंतर विशेष वर्गीकरण के आधार पर होता है इन्हें सामान्यत: दो वर्गों में वर्गीकृत किया जाता है:

(क) *संतत* तथा

(ख) *विविक्त*

संतत चर का कोई भी संख्यात्मक मान हो सकता है। यह पूर्णांक मान (1, 2, 3, 4 ...), भिन्नात्मक मान (1/2, 2/3, 3/4), तथा वे मान जो यथातथ भिन्न नहीं हैं ($\sqrt{2}$ = 1.414, $\sqrt{3}$ = 1.732, ... , $\sqrt{7}$ = 2.645) हो सकते हैं। उदाहरण के लिए, मान लिजिए कि एक छात्र का कद 90–150 सेमी तक बढ़ता है, तो उसके कद के मान इसके बीच आने वाले सभी मान हो सकते हैं। यह संपूर्ण संख्या वाले मान को भी प्रकट कर सकता है, जैसे कि 90 सेमी, 100 सेमी, 108 सेमी, 150 सेमी। इसके साथ ही यह भिन्नात्मक मान जैसे 90.85 सेमी, 102.34 सेमी, 149.99 सेमी आदि भी हो सकते हैं, जो पूर्णांक नहीं हैं। इस प्रकार 'ऊँचाई' चर किसी भी कल्पित मान को अभिव्यक्त करने में सक्षम है और इसके मानों को अनन्त श्रेणियों में बाँटा जा सकता है। संतत चर के अन्य उदाहरण भार, समय तथा दूरी आदि हैं।

संतत चर के विपरीत *विविक्त चर* केवल निश्चित मान हो सकते हैं। इसके मान केवल परिमित 'उछाल' से बदलते हैं। यह उछाल एक मान से दूसरे मान के बीच होते हैं, परंतु इसके बीच में कोई मान नहीं आता है। उदाहरण के लिए, कोई चर जैसे 'किसी कक्षा में छात्रों की संख्या', भिन्न वर्गों के लिए उन मानों की कल्पना करता है, जिसमें केवल पूर्ण संख्याएँ हों। यह कोई भी भिन्नात्मक मान जैसे 0.5 नहीं हो सकता, क्योंकि 'एक छात्र का आधा' निरर्थक है। इस प्रकार से इसमें 25 एवं 26 के बीच का मान 25.5 नहीं हो सकता है। इसकी अपेक्षा इसका मान या तो 25 होगा या फिर 26। हम देखते हैं कि जब इसका मान 25 से 26 में बदलता है, तो इन दोनों के बीच के भिन्नों को इसमें नहीं लिया जाता है। लेकिन ऐसा नहीं सोचना चाहिए कि किसी विविक्त चर का मान भिन्न में नहीं हो सकता। मान लीजिए कि **X** एक चर है जिसमें 1/8, 1/16, 1/32, 1/64 ..., जैसे मान हैं तो क्या यह एक विविक्त चर है? हाँ, क्योंकि यद्यपि

x के मान भिन्नों में हो सकते हैं, तथापि ये दो सन्निकट भिन्नों के बीच नहीं हो सकते। यह 1/8 से 1/16 में और फिर 1/16 से 1/32 में 'बदलता' है, परंतु 1/8 से 1/16 के बीच या 1/16 से 1/32 के बीच के मान नहीं ले सकता।

क्रियात्मक गतिविधि

- निम्नलिखित चरों का संतत तथा विविक्त में वर्गीकरण करें:
  क्षेत्रफल, आयतन, ताप, पाँसे पर आने वाली संख्या, फसल-उपज, जनसंख्या, वर्षा, सड़क पर कारों की संख्या और आयु।

हमने पहले यह बताया है कि उदाहरण 4 में 100 छात्रों के गणित में प्राप्तांक का बारंबारता वितरण दिया गया है, जैसा कि सारणी 3.1 में दिखाया गया है। यह दिखाता है कि 100 छात्रों के अंकों को वर्गों में कैसे समूहित किया गया है। आपको आश्चर्य होगा कि हमने सारणी 3.1 के अपरिष्कृत आँकड़ों से इसे कैसे प्राप्त किया। लेकिन इस प्रश्न का समाधान प्रस्तुत करने से पहले आपका यह जानना आवश्यक है कि बारंबारता वितरण क्या होता है।

5. बारंबारता वितरण क्या है?

बारंबारता वितरण अपरिष्कृत आँकड़ों को एक मात्रात्मक चर में वर्गीकृत करने का एक सामान्य तरीका है। यह दिखाता है कि किसी चर के भिन्न मान (यहाँ छात्र द्वारा गणित में प्राप्तांक) विभिन्न वर्गों में, अपने अनुरूप वर्गों की बारंबारताओं के साथ कैसे वितरित किए जाते हैं। इस उदाहरण में हमारे पास प्राप्तांकों के 10 वर्ग हैं। 0–10, 10–20, ... 90–100। *वर्ग-बारंबारता* पद का अर्थ है एक विशेष वर्ग में मानों की संख्या। उदाहरण के लिए वर्ग 30–40 में सारणी 3.1 में प्राप्तांकों के 7 मान हैं। ये 30, 37, 34, 30, 35, 39, 32 हैं। इस प्रकार से वर्ग 30–40 की बारंबारता 7 हुई। पर शायद आपको आश्यर्च हो कि 40 का अंक जो अपरिष्कृत आँकड़ों में दो बार आया है, उसे 30–40 वर्ग में शामिल क्यों नहीं किया गया? अगर इसे 30–40 वर्ग की बारंबारता में शामिल किया जाता, तो ये 7 की अपेक्षा 9 होते। यह पहेली तब स्पष्ट हो जाएगी, जब आप इस अध्याय को पर्याप्त धैर्य के साथ सावधानी पूर्वक पढ़ेंगे। इसलिए पढ़ना जारी रखें। आपको स्वयं ही इसका उत्तर प्राप्त हो जाएगा।

बारंबारता वितरण सारणी में प्रत्येक वर्ग, *वर्ग सीमाओं* द्वारा घिरा होता है। वर्ग में ये सीमाएँ दो छोरों पर होती हैं। इसमें न्यूनतम मान को *निम्नवर्ग सीमा* तथा उच्चतम मान को *उच्च वर्ग सीमा* कहते हैं। उदाहरण के लिए वर्ग 60–70 में वर्ग सीमाएँ 60 एवं 70 हैं। इसकी निम्न वर्ग सीमा 60 और उच्च वर्ग सीमा 70 है। *वर्ग मध्यांतर* या अंतराल या *वर्ग विस्तार* उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के बीच का अंतर है। वर्ग 60–70 के लिए वर्ग अंतराल 10 है, (उच्च वर्ग सीमा में से निम्न वर्ग सीमा को *घटाकर*)।

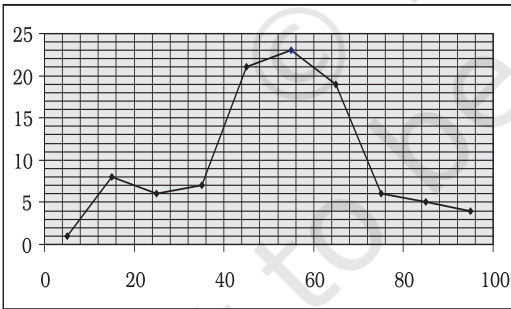
वर्ग मध्यबिन्दु अथवा *वर्ग चिह्न* किसी वर्ग का मध्य-मान है। यह वर्ग की निम्न वर्ग सीमा तथा उच्च वर्ग सीमा के बीच होता है। इसे निम्नलिखित तरीके से पता किया जा सकता है:

वर्ग मध्य बिन्दु या *वर्ग चिह्न = (उच्च वर्ग सीमा + निम्न वर्ग सीमा)/2 ... ... (1)*

प्रत्येक वर्ग का वर्ग चिह्न या वर्ग मध्य-बिन्दु एक वर्ग के प्रतिनिधित्व के लिए प्रयुक्त किया जाता है। एक बार जब अपरिष्कृत आँकड़ों को वर्गों में समूहित कर दिया जाता है, तब आगे की गणनाओं में व्यष्टि प्रेक्षणों का प्रयोग नहीं किए जाता है बल्कि इसकी जगह वर्ग चिह्न प्रयुक्त किया जाता है।

सारणी 3.3
निम्न वर्ग सीमा, उच्च वर्ग सीमा तथा वर्ग चिह्न

वर्ग	*बारंबारता*	*निम्नवर्ग सीमा*	*उच्चवर्ग सीमा*	*वर्ग चिह्न*
0–10	1	0	10	5
10–20	8	10	20	15
20–30	6	20	30	25
30–40	7	30	40	35
40–50	21	40	50	45
50–60	23	50	60	55
60–70	19	60	70	65
70–80	6	70	80	75
80–90	5	80	90	85
90–100	4	90	100	95

बारंबारता वक्र किसी बारंबारता वितरण का आलेखीय प्रस्तुतीकरण है। चित्र 3.1 के अंतर्गत उपरोक्त उदाहरण में दिए गए आँकड़ों का आरेखी प्रस्तुतीकरण दिया गया है। बारंबारता वक्र प्राप्त करने के लिए, हम वर्ग चिह्न को एक्स (x) अक्ष पर तथा बारंबारता को वाई (y) अक्ष पर आलेखित करते हैं।

चित्र 3.1 *आँकड़ों के बारंबारता वितरण का आरेखी प्रस्तुतीकरण।*

बारंबारता वितरण कैसे तैयार करें?

बारंबारता वितरण तैयार करते समय हमें निम्न पाँच प्रश्नों की व्याख्या पर ध्यान देने की आवश्यकता है:

1. वर्ग अंतराल समान आकार के हों या असमान आकार के?
2. हमें कितने वर्ग रखने चाहिए?
3. प्रत्येक वर्ग का आकार क्या हो?
4. वर्ग सीमाओं का निर्धारण कैसे किया जाय?
5. प्रत्येक वर्ग के लिए बारंबारता कैसे प्राप्त की जाय?

वर्ग अंतराल, समान अंतराल के हों या असमान अंतराल के?

दो परिस्थितियों में असमान आकार के वर्ग अंतरालों का प्रयोग किया जाता है। पहली, जब हमारे पास आय तथा ऐसे ही चरों के आँकड़े हों, जहाँ परास काफी अधिक होता है। उदाहरण के लिए, दैनिक आय लगभग शून्य से लेकर कई सौ करोड़ रुपये तक हो सकती है। ऐसी स्थिति में, समान वर्ग अंतराल उपयुक्त नहीं है, क्योंकि (i) यदि वर्ग अंतराल छोटे तथा समान आकार के होंगे, तो वर्गों की संख्या बहुत अधिक हो जाएगी। (ii) यदि वर्ग अंतराल अधिक है, तो आय के बहुत कम या बहुत अधिक स्तरों पर जानकारी छिपी हुई रह जाएगी।

दूसरी, यदि मानों की एक बहुत बड़ी संख्या परास के एक छोटे से भाग में केंद्रित होती है, तो समान वर्ग अंतराल से कई मानों की सूचना प्राप्त नहीं हो पाएगी।

अन्य सभी स्थितियों में, आवृत्ति-वितरण में समान आकार के वर्ग अंतरालों का प्रयोग होता है।

वर्गों की संख्या कितनी होनी चाहिए?

वर्गों की संख्या सामान्यत: 6 तथा 15 के बीच होती है। यदि हमारे वर्ग अंतराल समान आकार के हों, तो वर्गों की संख्या, परास (चर के अधिकतम तथा न्यूनतम मान में अंतर) को वर्ग अंतराल से भाग देने पर प्राप्त की जा सकती है।

क्रियात्मक गतिविधियाँ

निम्नलिखित का परास ज्ञात करें:
- उदाहरण 1 में भारत की जनसंख्या।
- उदाहरण 2 में गेहूँ की उपज।

प्रत्येक वर्ग का आकार क्या होना चाहिए?

इस प्रश्न का उत्तर पहले के प्रश्न के उत्तर पर निर्भर करता है। समीकरण (2) प्रकट करती है कि एक बार वर्ग अंतराल को तय करने पर चर के दिए गए परास से हम वर्गों की संख्या निर्धारित कर सकते हैं। ठीक इसी प्रकार से हम वर्ग अंतराल निर्धारित कर सकते हैं, जब एक बार हम वर्गों की संख्या तय कर लेते हैं। इस तरह हम पाते हैं कि ये दोनों निर्णय एक दूसरे से जुड़े हुए हैं। पहले का निर्णय लिए बिना हम दूसरे पर निर्णय नहीं ले सकते।

उदाहरण 4 में, हमारे पास वर्गों की संख्या 10 है तथा परास का दिया गया मान 100 है, तब वर्ग-अंतराल स्वत: ही (समानता 2 के द्वारा) 10 है। ध्यान दें कि वर्तमान संदर्भ में हमने वह वर्ग अंतराल चुना है, जिनका परिमाण समान है। तथापि हम ऐसा वर्ग अंतराल चुन सकते हैं जिसका परिमाण समान न हो, तब ऐसे मामले में वर्गों की चौड़ाई असमान होगी।

हमें वर्ग सीमाएँ कैसे निर्धारित करनी चाहिए?

वर्ग सीमाएँ निश्चित तथा स्पष्ट रूप से होनी चाहिए। सामान्ययत: मुक्तोत्तर वर्ग, जैसे– '70 तथा अधिक' या '10 से कम' वांछनीय नहीं होते। निम्न तथा उच्च वर्ग सीमाओं का निर्धारण इस प्रकार से किया जाना चाहिए कि प्रत्येक वर्ग की आवृत्तियों की प्रवृत्ति वर्ग अंतराल के मध्य में संकेंद्रण की हो। वर्ग अंतराल दो प्रकार के होते हैं–

1. समावेशी वर्ग अंतराल: इस स्थिति में, वर्ग की निम्न तथा उच्च सीमाओं के मूल्य वाले मानों को उस वर्ग की आवृत्ति में शामिल किया जाता है।
2. अपवर्जी वर्ग अंतराल: इस स्थिति में, वर्ग की निम्न तथा उच्च सीमाओं के मूल्य वाली मदों को उस वर्ग की आवृत्ति में शामिल नहीं किया जाता।

असतत चरों की स्थिति में, अपवर्जी तथा समावेशी, दोनों प्रकार के वर्ग अंतरालों का प्रयोग किया जा सकता है।

सतत चरों की स्थिति में, समावेशी वर्ग अंतरालों का प्रयोग बहुधा किया जाता है।

उदाहरण

मान लीजिए, हमारे पास एक परीक्षा में विद्यार्थियों द्वारा प्राप्तांकों के आँकड़े हैं तथा सभी प्राप्तांक पूर्णांक हैं (भिन्नात्मक अंकों की अनुमति नहीं है)। मान लीजिए, विद्यार्थियों द्वारा प्राप्तांक 0 से 100 के बीच हैं।

यह असतत चरों की स्थिति है, क्योंकि भिन्नात्मक अंकों की अनुमति नहीं है। इस स्थिति में, यदि हम समान आकार वाले वर्ग अंतरालों का उपयोग करते हैं तथा 10 वर्ग अंतरालों का प्रयोग करते हैं, तो वर्ग अंतरालों के निम्न रूप हो सकते हैं–

वर्ग अंतराल का समावेशी रूप

0–10
11–20
21–30
–
–
91–100

वर्ग अंतराल का अपवर्जी रूप

0–10
10–20
20–30
–
–
90–100

अपवर्जी वर्ग अंतराल की स्थिति में, हमें यह अग्रिम रूप से निर्धारित करना होता है कि वर्ग सीमा

के मान के बराबर किसी चर का मान होने पर क्या करना है। उदाहरण के लिए, हम यह निर्णय कर सकते हैं कि 10, 30 आदि मानों को क्रमशः वर्ग अंतराल ''0 से 10'' तथा ''20 से 30'' में रखा जाए। इस स्थिति में वर्ग की निचली सीमा को वर्ग अंतराल में शामिल नहीं किया जाता।

या फिर हम 10, 30 आदि मानों को क्रमशः वर्ग अंतराल ''10 से 20'' तथा ''30 से 40'' में रख सकते हैं। इस स्थिति में वर्ग की उच्च सीमा को वर्ग अंतराल में शामिल नहीं किया जाता।

सतत चर के उदाहरण

मान लें कि हमारे पास किसी चर के आँकड़े उपलब्ध हों, जैसे कद (से.मी.) या वज़न (कि.ग्रा.)। यह आँकड़ा सतत प्रकार का है। ऐसी स्थितियों में वर्ग अंतराल निम्नलिखित प्रकार से दर्शाया जा सकता है–

30 कि.ग्रा.– 39.999...कि.ग्रा.
40 कि.ग्रा.– 49.999...कि.ग्रा.
50 कि.ग्रा.– 59.999...कि.ग्रा. आदि।

इन वर्ग अंतरालों को निम्नलिखित प्रकार से समझा जा सकता है–

30 कि.ग्रा. और अधिक तथा 40 कि.ग्रा. से कम
40 कि.ग्रा. और अधिक तथा 50 कि.ग्रा. से कम
50 कि.ग्रा. और अधिक तथा 60 कि.ग्रा. से कम आदि।

सारणी 3.4
एक कंपनी के 550 कर्मचारियों की आय का बारंबारता वितरण

आय (रु में)	*कर्मचारियों की संख्या*
800–899	50
900–999	100
1000–1099	200
1100–1199	150
1200–1299	40
1300–1399	10
योग	*550*

वर्ग अंतराल में समायोजन

सारणी 3.4 में समावेशी विधि के सूक्ष्म अध्ययन से पता चलता है कि यद्यपि चर 'आय' एक संतत चर है, तथापि जब वर्गों को बनाया जाता है तो संततता नहीं रहती। हम एक वर्ग की उच्च सीमा तथा अगले वर्ग की निम्न सीमा में 'अंतर' या असंततता पाते हैं। उदाहरण के लिए, पहले वर्ग की उच्च सीमा 899 और दूसरे वर्ग की निम्न सीमा 900 के बीच हम 1 (एक) का 'अंतर' पाते हैं। तब हम आँकड़ों के वर्गीकरण में चर की संततता को कैसे सुनिश्चित करते हैं? इसे वर्ग अंतराल के बीच समायोजन करके किया जाता है। समायोजन निम्नलिखित तरीके से किया गया है।

1. द्वितीय वर्ग की निम्न सीमा और प्रथम वर्ग की उच्च सीमा के बीच अंतर पता करें। उदाहरण के लिए, सारणी 3.4 में द्वितीय वर्ग की निम्न सीमा 900 और प्रथम वर्ग की उच्च सीमा 899 के बीच अंतर 1 है (अर्थात 900 – 899 = 1)।
2. प्राप्त किए गए अंतर (1) को 2 से विभाजित करें (अर्थात 1/2 = 0.5)।
3. सभी वर्गों की निम्न सीमाओं से (2) में प्राप्त किए गए मान को घटाइए (निम्न वर्ग सीमा – 0.5)।
4. सभी वर्गों की उच्च सीमा में (2) में प्राप्त किए गए मान को जोड़िए (उच्च वर्ग सीमा + 0.5)।

समायोजन के पश्चात्, जिससे बारंबारता वितरण में आँकड़ों की संततता की पुनः प्राप्ति होती है, सारणी 3.4 संशोधित होकर सारणी 3.5 बन जाती है।

वर्ग सीमाओं में समायोजन के पश्चात्, समानता (1) जोकि वर्ग चिह्न का मान निर्धारित करती है, निम्नलिखित प्रकार से संशोधित हो जाएगी:

समायोजित वर्ग चिह्न = (समायोजित उच्च वर्ग सीमा + समायोजित निम्न वर्ग सीमा)/2

सारिणी 3.5

एक कंपनी के 550 कर्मचारियों की आय का बारंबारता वितरण

आय (रु में)	*कर्मचारियों की संख्या*
799.5–899.5	50
899.5–999.5	100
999.5–1099.5	200
1099.5–1199.5	150
1199.5–1299.5	40
1299.5–1399.5	10
योग	*550*

हमें प्रत्येक वर्ग की बारंबारता कैसे प्राप्त करनी चाहिए

साधारण शब्दों में, *एक प्रेक्षण की बारंबारता का अर्थ है कि अपरिष्कृत आँकड़ों में कितनी बार वह प्रेक्षण प्रकट होता है।* सारणी 3.1 में, हमने देखा कि 40 का मान तीन बार आया है, जबकि 0 और 10 का मान एक बार, 49 का मान 5 बार और ऐसे ही अन्य मान आये हैं। इस प्रकार से 40 की बारंबारता 3, 0 की 1, 10 की 1, 49 की 5 तथा ऐसे ही। लेकिन जब आँकड़े वर्गों में समूहित कर दिए जाते हैं, जैसा कि उदाहरण 3 में किया गया है, तो किसी वर्ग की बारंबारता से तात्पर्य उस वर्ग के मानों की संख्याओं से है। वर्ग-बारंबारताओं की गिनती विशेष वर्ग के सामने मिलान चिह्नों को लगाकर की जाती है।

मिलान चिह्न अंकन द्वारा वर्ग बारंबारता को ज्ञात करना

मिलान चिह्न (/) किसी वर्ग के प्रत्येक छात्र के सामने लगाया जाता है, जिसके प्राप्तांक उस वर्ग में शामिल हैं। उदाहरण के लिए, यदि किसी छात्र का प्राप्तांक 57 है तो उस छात्र के लिए वर्ग 50–60 में एक मिलान चिह्न (/) लगाया जाता है। यदि प्राप्तांक 71 हैं तो मिलान चिह्न (/) को वर्ग 70–80 में लगाया जाता है। यदि कोई 40 अंक प्राप्त करता है तो उसके लिए मिलान चिह्न वर्ग 40–50 में लगाया जाता है। सारणी 3.1 के 100 छात्रों के गणित में प्राप्तांकों के मिलान चिह्नों को सारणी 3.6 में दिखाया गया है।

मिलान चिह्नों का परिकलन तब आसान हो जाता है जब 4 चिह्न खड़े (////) लगाए जाते हैं और पाँचवाँ चिह्न सबको काटता हुआ तिरछा लगाया जाता है, जैसे (/~~///~~/)। मिलान चिह्नों की गणना पाँच के समूह में की जाती है, इसलिए यदि किसी वर्ग में 16 मिलान चिह्न हैं तो उन्हें इस प्रकार से /~~///~~/ /~~///~~/ /~~///~~/ / लिखते हैं, ताकि परिकलन में सुविधा रहे। इसलिए एक वर्ग की बारंबारता उतनी ही होगी, जितनी उस वर्ग में मिलान चिह्नों की संख्या।

सूचना की हानि (Loss of Information)

बारंबारता वितरण के रूप में आँकड़ों के वर्गीकरण में एक अंतर्निहित दोष पाया जाता है। यह अपरिष्कृत आँकड़ों का सारांश प्रस्तुत कर उन्हें संक्षिप्त एवं बोध गम्य तो बनाता है, परंतु इसमें वे विस्तृत विवरण नहीं प्रकट हो पाते जो अपरिष्कृत आँकड़ों में पाए जाते हैं। यद्यपि अपरिष्कृत आँकड़ों को वर्गीकृत करने में सूचना की क्षति होती है, तथापि आँकड़ों को वर्गीकरण द्वारा संक्षिप्त करने पर पर्याप्त जानकारी मिल जाती है। एक बार जब आँकड़ों को वर्गों में समूहित कर दिया जाता है तब व्यष्टि प्रेक्षणों का आगे सांख्यकीय परिकलनों में कोई महत्व नहीं होता। उदाहरण 4 में, वर्ग 20–30 के अंतर्गत 6 प्रेक्षण 25, 25, 20, 22, 25 एवं 28 हैं। इसलिए जब इन आँकड़ों को बारंबारता वितरण में वर्ग 20–30 में समूहित कर दिया जाता है, तब यह बारंबारता वितरण उस वर्ग की बारंबारता (जैसे 6) को दिखाता है, न कि उनके वास्तविक मानों को। *इस वर्ग के सभी मानों को उस वर्ग के वर्ग-अंतराल के मध्य मान या वर्ग चिह्न के बराबर माना जाता है (अर्थात् 25)। आगे की सांख्यिकीय परिकलनों के लिए वर्ग चिह्न के मान को आधार बनाया जाता है, न*

कि उस वर्ग के प्रेक्षणों के मान को। यही बात सभी वर्गों के लिए सत्य है। इस प्रकार प्रेक्षणों के वास्तविक मान के स्थान पर वर्ग चिह्नों के प्रयोग को सांख्यिकीय विधियों में शामिल करने पर पर्याप्त मात्रा में सूचनाओं की क्षति होती है। यद्यपि अपरिष्कृत आँकड़ों का जैसा कि नीचे इससे अधिक दिखाया गया है, अधिक अर्थपूर्ण लगता है।

असमान वर्गों में बारंबारता वितरण

अब तक आप समान वर्ग अंतराल के बारंबारता वितरण से परिचित हो चुके हैं। आप जान गए हैं कि इन्हें अपरिष्कृत आँकड़ों से कैसे गठित किया जाता है। लेकिन कुछ मामलों में असमान वर्ग अंतराल के साथ बारंबारता वितरण अधिक उपयुक्त होता है। यदि आप उदाहरण 4 के बारंबारता वितरण की *सारणी 3.6* को देखें, तो आप पायेंगे कि अधिकांश प्रेक्षण वर्ग 40–50, 50–60 तथा 60–70 में संकेंद्रित हैं। उनकी बारंबारताएँ क्रमशः 21, 23 एवं 19 हैं। इसका अर्थ है कि 100 छात्रों में से 63 (21 + 23 + 19) प्रेक्षण इन वर्गो में सकेंद्रित हैं। इस प्रकार 63 प्रतिशत आँकड़े 40–70 के बीच समाहित है और आँकड़ों का शेष 37 प्रतिशत 0–10, 10–20, 20–30, 30–40 तथा 70–80, 80–90 एवं 90–100 वर्गों में हैं। इन वर्गों में प्रेक्षणों का विरल घनत्व है। आप यह भी देखेंगे कि इन वर्गों के प्रेक्षणों में अन्य वर्गों की अपेक्षा उनके अपने वर्गों के वर्ग-चिह्नों से अधिक विचलन है। लेकिन यदि वर्गों का गठन इस प्रकार से करना हो कि वर्ग चिह्न, जहाँ तक संभव है, उस मान के बराबर हो जाए, जिसके आस-पास उस वर्ग के प्रेक्षणों के संकेंद्रण की प्रवृत्ति होती है, तो असमान वर्ग अंतराल अधिक उपयुक्त होता है।

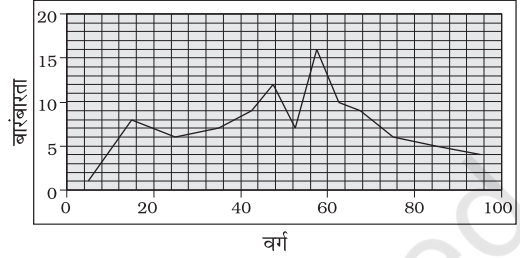
असमान वर्गों के रूप में, सारणी 3.7 में सारणी 3.6 के उसी बारंबारता वितरण को दिखाया गया है। 40–50, 50–60 तथा 60–70 के प्रत्येक वर्ग को दो भागों में विभाजित किया गया है। वर्ग 40–50 को

सारणी 3.6

गणित में 100 छात्रों के प्राप्तांको के मिलान चिह्न

वर्ग	*प्रेक्षण*	*मिलान चिह्न*	*बारंबारता*	*वर्ग चिह्न*
0–10	0	/	1	5
10–20	10, 14, 17, 12, 14, 12, 14, 14	~~////~~ ///	8	15
20–30	25, 25, 20, 22, 25, 28	~~////~~ /	6	25
30–40	30, 37, 34, 39, 32, 30, 35,	~~////~~ //	7	35
40–50	47, 42, 49, 49, 45, 45, 47, 44, 40, 44, 49, 46, 41, 40, 43, 48, 48, 49, 49, 40, 41	~~////~~ ~~////~~ ~~////~~ ~~////~~ /	21	45
50–60	59, 51, 53, 56, 55, 57, 55, 51, 50, 56, 59, 56, 59, 57, 59, 55, 56, 51, 55, 56, 55, 50, 54	~~////~~ ~~////~~ ~~////~~ ~~////~~ ///	23	55
60–70	60, 64, 62, 66, 69, 64, 64, 60, 66, 69, 62, 61, 66, 60, 65, 62, 65, 66, 65	~~////~~ ~~////~~ ~~////~~ ////	19	65
70–80	70, 75, 70, 76, 70, 71	~~////~~ /	6	75
80–90	82, 82, 82, 80, 85	~~////~~	5	85
90–100	90, 100, 90, 90	////	4	95
	योग		*100*	

अब 40–45 तथा 45–50 में बाँटा गया है। वर्ग 50–60 को 50–55 और 55–60 में बाँटा गया है तथा वर्ग 60–70 को 60–65 तथा 65–70 में बाँटा गया है। अब नए वर्ग 40–45, 45–50, 50–55, 55–60, 60–65 तथा 65–70 हैं जिनमें वर्ग अंतराल 5 है। बाकी अन्य वर्गों 0–10, 10–20, 20–30, 30–40, तथा 70–80, 80–90, 90–100 में ठीक वही पूर्ववत वर्ग अंतराल 10 है। इस सारणी का अंतिम स्तंभ इन वर्गों के नये वर्ग चिह्नों को प्रदर्शित कर रहा है। सारणी 3.6 के पुराने वर्ग चिह्नों से उनकी तुलना करें। ध्यान दें कि इन वर्गों के प्रेक्षणों में नये वर्ग चिह्न मानों की अपेक्षा पुराने वर्ग चिह्न मानों से विचलन अधिक है। इस प्रकार से नये वर्ग चिह्न मान, इन वर्गों के आँकड़ों का पुराने मान की अपेक्षा बेहतर प्रतिनिधित्व करते हैं।

चित्र 3.2 में, सारणी 3.7 के बारंबारता वितरण के बारंबारता वक्र को दिखाया गया है। इसमें सारणी के वर्ग चिह्नों को x-अक्ष पर तथा बारंबारताओं को y-अक्ष पर आलेखित किया गया है।

3.2 बारंबारता वक्र

क्रियाकलाप

- यदि आप चित्र 3.2 के साथ चित्र 3.1 की तुलना करते हैं तो आप क्या देखते हैं? क्या आपने इनके बीच कोई अंतर पाया? क्या आप उस अंतर की व्याख्या कर सकते हैं?

सारणी 3.7
असमान वर्गों में बारंबारता वितरण

वर्ग	*प्रेक्षण*	*बारंबारता*	*वर्ग चिह्न*
0–10	0	1	5
10–20	10, 14, 17, 12, 14, 12, 14, 14	8	15
20–30	25, 25, 20, 22, 25, 28	6	25
30–40	30, 37, 34, 39, 32, 30, 35	7	35
40–45	42, 44, 40, 44, 41, 40, 43, 40, 41	9	42.5
45–50	47, 49, 49, 45, 45, 47, 49, 46, 48, 48, 49, 49	12	47.5
50–55	51, 53, 51, 50, 51, 50, 54	7	52.5
55–60	59, 56, 55, 57, 55, 56, 59, 56, 59, 57, 59, 55, 56, 55, 56, 55	16	57.5
60–65	60, 64, 62, 64, 64, 60, 62, 61, 60, 62	10	62.5
65–70	66, 69, 66, 69, 66, 65, 65, 66, 65	9	67.5
70–80	70, 75, 70, 76, 70, 71	6	75
80–90	82, 82, 82, 80, 85	5	85
90–100	90, 100, 90, 90	4	95
	योग	*100*	

बारंबारता सरणी (Frequency Array)

अब तक हमने गणित में 100 छात्रों द्वारा प्राप्त किए गए प्रतिशत अंकों के उदाहरण का प्रयोग करते हुए संतत चर के लिए आँकड़ों के वर्गीकरण पर चर्चा की है। विविक्त चर के लिए, आँकड़ों का वर्गीकरण *बारंबारता सरणी* के नाम से जाना जाता है। चूँकि एक विविक्त चर मानों को धारण करता है न कि दो पूर्णांकों के बीच माध्यमिक भिन्नीय मानों को, अतः हम ऐसी बारंबारता रखते हैं जोकि अपने पूर्णांक मानों से संगत हों।

सारणी 3.8
परिवारों के आकार की बारंबारता सारणी

परिवार का आकार	*परिवारों की संख्या*
1	5
2	15
3	25
4	35
5	10
6	5
7	3
8	2
योग	*100*

सारणी 3.8 में दिया गया उदाहरण बारंबारता सरणी को प्रदर्शित करता है। इस सारणी में चर 'परिवार का आकार' एक विविक्त चर है जो सारणी में दिखाए गए पूर्णांकों को ही धारण करता है।

6. द्विचर बारंबारता वितरण

बहुत बार, जब हम किसी जनसंख्या में से एक प्रतिदर्श लेते हैं, तो हम प्रतिदर्श के हर अवयव से एक से अधिक प्रकार की सूचना संगृहीत करते हैं। उदाहरण के लिए, मान लीजिए कि हमने एक शहर की कंपनियों की सूची में से 20 कंपनियों का एक प्रतिदर्श लिया है। मान लीजिए कि हम प्रत्येक कंपनी से P बिक्री तथा विज्ञापनों पर किए गए व्यय की जानकारी संगृहीत करते हैं। इस स्थिति में, हमारे पास प्रतिदर्श के द्विचर आँकड़े हैं। इस तरह के द्विचर आँकड़ों को द्विचर बारंबारता वितरण द्वारा संक्षिप्त रूप में दर्शाया जा सकता है।

एक द्विचर बारंबारता वितरण को दो चरों के बारंबारता वितरण के रूप में परिभाषित किया जा सकता है।

सारणी 3.9, 20 कंपनियों के दो चर-बिक्री एवं विज्ञापन व्यय (लाख रु में) के बारंबारता वितरण को प्रदर्शित कर रही है। यहाँ पर बिक्री मानों को भिन्न स्तंभों में तथा विज्ञापन व्यय के मानों को भिन्न पंक्तियों में वर्णित किया गया है। प्रत्येक प्रकोष्ठ संतत पंक्ति एवं स्तंभ के मान की बारंबारता दिखाता है। उदाहरण के लिए, यहाँ पर तीन फर्म ऐसी हैं, जिनकी बिक्री रु 135–145 लाख रु के

सारणी 3.9
20 कंपनियों की बिक्री (लाख रु में) एवं विज्ञापन व्यय (हजार रु में) का द्विचर बारंबारता वितरण

	115–125	*125–135*	*135–145*	*145–155*	*155–165*	*165–175*	*योग*
62–64	2	1					3
64–66	1		3				4
66–68	1	1	2	1			5
68–70		2		2			4
70–72		1	1		1	1	4
योग	*4*	*5*	*6*	*3*	*1*	*1*	*20*

बीच है और उनका विज्ञापन व्यय 64–66 हजार रु के बीच है। द्विचर वितरण के बारे में अध्याय 8 'सहसंबंध' में अध्ययन किया जाएगा।

7. सारांश

प्राथमिक या द्वितीयक स्रोतों से संगृहीत किए गए आँकड़े अपरिष्कृत या अवर्गीकृत होते हैं। जब एक बार आँकड़े संगृहित हो जाएँ तो अगला चरण आगे के सांख्यिकीय विश्लेषण के लिए आँकड़ों का वर्गीकरण करना है। वर्गीकरण से आँकड़ों में क्रमबद्धता आ जाती है। यह अध्याय आप को यह जानने के योग्य बनाता है कि आँकड़ों को बारंबारता वितरण के माध्यम से बोधगम्य तरीके से किस प्रकार वर्गीकृत किया जाता है। एक बार जब आप वर्गीकरण की तकनीकों को जान जाते हैं तो आपके लिए यह आसान होगा कि आप संतत तथा विविक्त दोनों चरों के लिए ही बारंबारता वितरण की रचना कर सकें।

पुनरावर्तन

- वर्गीकरण अपरिष्कृत आँकड़ों को क्रमबद्धता प्रदान करता है।
- बारंबारता वितरण यह प्रदर्शित करता है कि किसी चर के विभिन्न मान, संगत वर्ग-बारंबारताओं सहित, किस प्रकार विभिन्न वर्गों में वितरित किए जाते हैं।
- अपवर्जी विधि के अंतर्गत उच्च वर्ग सीमा को छोड़ा तथा निम्नवर्ग सीमा को शामिल किया जाता है।
- समावेशी विधि में निम्नवर्ग सीमा तथा उच्च वर्ग सीमा, दोनों को ही शामिल किया जाता है।
- बारंबारता वितरण में, आगे के सांख्यिकीय परिकलन केवल वर्ग चिह्न मान पर आधारित होते हैं, न कि प्रेक्षणों के मान पर।
- वर्गों को इस प्रकार से बनाया जाना चाहिए कि, जहाँ तक संभव हो सके, प्रत्येक वर्ग का वर्ग चिह्न उस मान के अधिक से अधिक निकटतम हो, जिस मान के आस-पास, किसी वर्ग के प्रेक्षणों की संकेन्द्रण की प्रवृत्ति हो।

अभ्यास

1. निम्नलिखित में से कौन सा विकल्प सही है?
   - एक वर्ग मध्यबिन्दु बराबर है:
     (क) उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के औसत के।
     (ख) उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के गुणनफल के।
     (ग) उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के अनुपात के।
     (घ) उपरोक्त में से कोई नहीं।
   - दो चरों के बारंबारता वितरण को इस नाम से जानते हैं:
     (क) एक विचर वितरण
     (ख) द्विचर वितरण
     (ग) बहुचर वितरण
     (घ) उपरोक्त में से कोई नहीं
   - वर्गीकृत आँकड़ों में साँख्यिकीय परिकलन आधारित होता है:
     (क) प्रेक्षणों के वास्तविक मानों पर
     (ख) उच्च वर्ग सीमाओं पर
     (ग) निम्न वर्ग सीमाओं पर
     (घ) वर्ग के मध्यबिन्दुओं पर
   - अपवर्जी विधि के अंतर्गत
     (क) किसी वर्ग की उच्च वर्ग सीमा को वर्ग अंतराल में समावेशित नहीं करते।
     (ख) किसी वर्ग की उच्च वर्ग सीमा को वर्ग अंतराल में समावेशित करते हैं।
     (ग) किसी वर्ग की निम्न वर्ग सीमा को वर्ग अंतराल में समावेशित नहीं करते।
     (घ) किसी वर्ग की निम्न वर्ग सीमा को वर्ग अंतराल में समावेशित करते हैं।
   - परास का अर्थ है
     (क) अधिकतम एवं न्यूनतम प्रेक्षणों के बीच अंतर
     (ख) न्यूनतम एवं अधिकतम प्रेक्षणों के बीच अंतर
     (ग) अधिकतम एवं न्यूनतम प्रेक्षणों का औसत
     (घ) अधिकतम एवं न्यूनतम प्रेक्षणों का अनुपात
2. वस्तुओं को वर्गीकृत करने में क्या कोई लाभ हो सकता है? अपने दैनिक जीवन से एक उदाहरण देकर व्याख्या कीजिए।
3. चर क्या है? एक संतत तथा विविक्त चर के बीच भेद कीजिए।
4. आँकड़ों के वर्गीकरण में प्रयुक्त अपवर्जी तथा समावेशी विधियों की व्याख्या कीजिए।

5. सारणी 3.2 के आँकड़ों का प्रयोग करें, जो 50 परिवारों के भोजन पर मासिक व्यय (रु में) को दिखलाती है, और
   (क) भोजन पर मासिक परिवारिक व्यय का प्रसार ज्ञात कीजिए।
   (ख) परास को वर्ग अंतराल की उचित संख्याओं में विभाजित करें तथा व्यय का बारंबारता वितरण प्राप्त करें।
   - उन परिवारों की संख्या पता कीजिए जिनका भोजन पर मासिक व्यय
     (क) 2000/- रु से कम है
     (ख) 3000/- रु से अधिक है
     (ग) 1500/- रु और 2500/- रु के बीच है

6. एक शहर में, यह जानने हेतु 45 परिवारों का सर्वेक्षण किया गया कि वे अपने घरों में कितनी संख्या में सेल फोनों का इस्तेमाल करते हैं। नीचे दिए गए उनके उत्तरों के आधार पर एक बारंबारता सरणी तैयार कीजिए।

1	3	2	2	2	2	1	2	1	2	2	3	3	3	3
3	3	2	3	2	2	6	1	6	2	1	5	1	5	3
2	4	2	7	4	2	4	3	4	2	0	3	1	4	3

7. वर्गीकृत आँकड़ों में 'सूचना की क्षति' का क्या अर्थ है?

8. क्या आप इस बात से सहमत है कि अपरिष्कृत आँकड़ों की अपेक्षा वर्गीकृत आँकड़े बेहतर होते हैं?

9. एक-विचर एवं द्विचर बारंबारता वितरण के बीच अंतर बताइए?

10. निम्नलिखित आँकडों के आधार पर 7 का वर्ग अंतराल लेकर समावेशी विधि द्वारा एक बारंबारता वितरण तैयार कीजिए।

28	17	15	22	29	21	23	27	18	12	7	2	9	4
1	8	3	10	5	20	16	12	8	4	33	27	21	15
3	36	27	18	9	2	4	6	32	31	29	18	14	13
15	11	9	7	1	5	37	32	28	26	24	20	19	25
19	20	6	9										

क्रियात्मक गतिविधि

- अपनी पुरानी अंक सारणियों से, अपनी पूर्व कक्षा में अर्द्धवार्षिक तथा वार्षिक परीक्षाओं में प्राप्त गणित के प्राप्तांकों को पता कीजिए। इन्हें वर्ष के क्रम में व्यवस्थित कीजिए। अब यह जाँच कीजिए कि क्या उक्त विषय में आप द्वारा प्राप्त किए अंक चर हैं या नहीं। इसके साथ यह भी देखिए कि क्या बाद के वर्षों में गणित में आपकी स्थिति में सुधार हुआ है?